

How Communities Are Exploited On Our Platforms: A Final Look At The “Troll Farm” Pages

JEFF ALLEN (D4TA SCIENCE) · FRIDAY, OCTOBER 4, 2019 · READING TIME: 28 MINUTES

My last note! Only fitting that I end where I started in integrity. Here is my summary and takeaways from a year and a half studying the Troll Farms.

Background

In 2016, a group of Pages burst forth into the American political scene. The owners of the Pages did not speak English particularly well, or Spanish, or any other language commonly spoken in America. They did not have any real understanding of American politics. And their primary interest in entering the American political publishing scene was to make a fast buck and cause chaos. Regardless, they were able to become a significant fraction of all civic discourse on our platform leading into the 2016 election.

It is profoundly strange that people who do not speak the local language, do not understand US politics, and fundamentally do not care about US politics, were able to become one of the loudest voices in the civic space on FB.

In the three years since, we have created a large policy and monitoring apparatus that helps us squash these actors when they pop up in the civic space. And we have created guardrail demotions that prevent the worst of the worst. But we have not fundamentally changed the way our platform operates to make it resilient against these actors.

The “Troll Farms” are a loosely defined set of FB Pages run out of the Balkans that target foreign audiences. Primarily US, but they also target Central and South American countries, UK and Australia, and India.

They have been a consistent problem and have been regularly flagged in the press from 2016 until just this week. They have been discussed by a number of news publishers and led to a significant amount of negative press for our company ([The Washington Post](#), [Newsweek](#), [CNN](#), [Wired](#), [BBC](#), [ABC News Australia](#), and [Radio Free Europe](#)).

There are around 15,000 FB Pages run out of Kosovo and Macedonia with a majority US audience (The countries where these behaviors are most popular, but the behavior has spread to neighboring countries as well, such as Slovenia and Albania [REDACTED], “[Low Quality Civic Exports Targeting the US](#)”). We typically treat each Page individually when studying them. But in this analysis, I would like to treat them as one single Page, since there is such a small set of coordinated actors behind all of them.

They also talk about only a few subjects. Primarily: targeting religious groups, targeting African Americans, targeting women. I will also group their topic Pages into one Page in this analysis.

Doing this will allow us to understand to total impact these Troll Farm Pages are having, relative to authentic Pages within those communities.

By the numbers

From: <https://fburl.com/unidash/sfsydwlo> and <https://fburl.com/unidash/jyaxzh9t>

1. They reach 140M US users monthly
 1. 0.75% of US, Pages VPVs
 2. 100M US users weekly
 3. 360M global users weekly
2. When combined, *the Troll Farm Pages form the largest Page on FB, by 40%*
 1. Reaching 140M US users monthly
 1. 75% of their reach is non-direct, meaning, the majority of users who see their content never followed the Page
 2. 2nd place Page (Walmart) reaches 100M US users monthly
3. When combined, *the Troll Farms run the largest Christian American Page on FB, by 20x*
 1. Reach 75M US users monthly
 1. 95% of their reach is non-direct
 2. Largest non-Troll Farm Christian American Page on FB (Guideposts) reaches 3M
 1. There are 16 Troll Farm Pages bigger than Guideposts
4. When combined, *the Troll Farms run the largest African American Page on FB, by 3x*
 1. Reach 30M US users monthly
 1. 85% of their reach is non-direct
 2. Largest non-Troll Farm African American Page on FB (Essence) reaches 12M
5. When combined, *the Troll Farms run the 2nd largest Native American Page on FB*
 1. Reach 400k US users monthly
 1. 90% of their reach is non-direct
 2. Largest Native American Page (Indian Country Today) reached 500k
 1. So comparable
6. When combined, *the Troll Farms run the 5th largest Womens Page on FB*
 1. Reach 60M US users monthly
 1. 90% of their reach is non-direct
 2. Largest Womens Page (Blossom) reaches 80M
 1. So comparable

So What?

Pages is the broadcast distribution product on our platform. It enables individuals or organizations to reach large audiences. It is where the “largest voices” of our communities live. It is how community leaders can communicate and interact with community members. It is how new community leaders are formed. It is such a critical tool to help us accomplish our company mission.

Our platform has given the largest voice in the Christian American community to a handful of bad actors, who, based on their media production practices, have never been to church.

Our platform has given the largest voice in the African American community to a handful of bad actors, who, based on their media production practices, have never had an interaction with an African American.

Our platform has given the 2nd largest voice in the Native American community to a handful of bad actors, who, based on their media production practices, have never had an interaction with a Native American.

Our platform has given the 5th largest voice in the American Womens community to a handful of bad actors, who, based on their media production practices, are not women and don't particularly care about any issues confronting women.

This is not normal. This is not healthy. We have empowered inauthentic actors to accumulate huge followings for largely unknown purposes. Mostly, they seem to want to skim a quick buck off of their audience. But there are signs they have been in [contact with the IRA](#). The fact that actors with possible ties to the IRA have access to huge audience numbers in the same [demographic groups targeted by the IRA](#) poses an enormous risk to the US 2020 election.

In addition, the behaviors and tactics of the Troll Farms are identical to the large broadcast tactics the IRA used in 2016, the IT Cells used in 2019, and misinfo posters have used time and again. As long as Troll Farms are able to find success on our platforms, that means that the IRA, IT Cells, and any other foreign government operation will be able to find success in the broadcast distribution space. If the Troll Farms are reaching 30M US users with content targeted to African Americans, we should not at all be surprised if we discover the IRA also currently has large audiences there.

This Note

In this Note, I'll cover how the Pages operate, how our systems empower them, why they do it, how our policies empower them, how many similar actors are out there, and what we could do to reduce their success.

What Do They Look Like?

You can find the full list here: <https://fburl.com/unidash/edgb1tk6>

African American Pages

Here are the top Troll Farm Pages targeting the African American community on FB, ranked by reach. The left is purely the Troll Farm Pages, the right is a mixture of the Troll Farm Pages plus the top authentic Pages. Top authentic Pages were sourced from [here](#), [here](#), and [here](#).

Ranked Troll Farm Pages				Ranked Troll Farm + Authentic Pages		
name	is_mkxxk_page	page_id		name	is_mkxxk_page	page_id
OVERALL_MKXXK	1	0		OVERALL_MKXXK	1	0
My Baby Daddy Aint Shit	1	113350735396543		My Baby Daddy Aint Shit	1	113350735396543
Hair & Beauty	1	248328492542716		Essence	0	68771922854
BlackBabies	1	318128722170152		The Root	0	26423400230
Gang Memes Videos	1	1160926057444860		The Source Magazine	0	201170940075
Savage Hood	1	191649061530625		BlackBabies	1	318128722170
Hood Life Memes	1	443152743162647		Hair & Beauty	1	248328492542
Savage Hood Memes	1	223612214784598		Gang Memes Videos	1	116092605744
Hood Mood Memes	1	672418553189918		Savage Hood	1	191649061530
HooDope	1	2350545161841610		Hood Life Memes	1	443152743162
Ghetto Memes	1	180345115786018		Black Enterprise	0	16037737277
The Positive Black Images	1	143354602735614		Savage Hood Memes	1	223612214784
Hood Videos	1	857393307661298		Hood Mood Memes	1	672418553189
Texas Smile	1	298117550830647		HELLO BEAUTIFUL	0	51757123845
Funny Hood Videos	1	415860738747398		HooDope	1	235054516184
BlackWomen	1	248882492463202		The Positive Black Images	1	143354602735

On the left: The top Troll Farm Pages built around content from African Americans, ranked by reach. On the right: Similar to the left, but the top authentic African American Pages are included for comparison.

6 out of the top 10 Pages with content from the African American community are Troll Farm Pages. Even the top individual Troll Farm Page outranks serious, well known, and valued African American outlets.

The content the Troll Farms post wouldn't be classified as violating, or even borderline. A lot of it is pretty entertaining. The babies are definitely cute. But it is a sign our platform is malfunctioning that a significant reward for the production of this content is going to known bad actors who are taking this content from actual African American creators. And the whole point of these Pages is to accumulate as large an audience as possible strictly for financial, and potentially political, reasons.



Example posts from Troll Farm Pages built around African American content.

They source the content primarily from Twitter and IG accounts. They likely found a pool of accounts to follow, and they post the most popular content from them.

There are also signs that the admins are successfully fooling our location transparency product here. The BlackBabies Page reports it has 1 admin in the United States. But the integrity focused location prediction systems place all admins in Kosovo [REDACTED], "Discrepancies Between Home Prediction And Zero Trust Prediction".

Religious Pages

Here are the top Troll Farm Pages targeting religious communities on FB. They primarily focus on the Christian community. But they do have a few Pages targeting Muslims. The top authentic Pages were sourced from [here](#) and [here](#).

Ranked Troll Farm Pages

name	is_mkxx_page	page_id
OVERALL_MKXX	1	0
Be Happy Enjoy Life	1	249346019035747
Jesus is my Lord	1	944383662417356
You Can	1	746720989059699
Speak Loudly	1	337270806881846
Purpose of Life	1	175964166645416
Positive Quotes	1	1537484389716090
Memories of a loved one	1	457990627606884
Smile & Shine	1	658641567851336
Light of the World	1	171049673814368
Why Not Us?	1	190698914861953
Happiness Is In Your Mind	1	1974991652809180
My Life	1	749005765476444
Life is Beautiful	1	277543839631924
Blessing	1	291773748400681
Conscious Reminder	1	160340534382111
Blessed	1	437169703776552
Simply Beautiful	1	272199883444219
Motivational Quotes & Videos	1	373989519691000
light-candle for a lost love one	1	45298736196

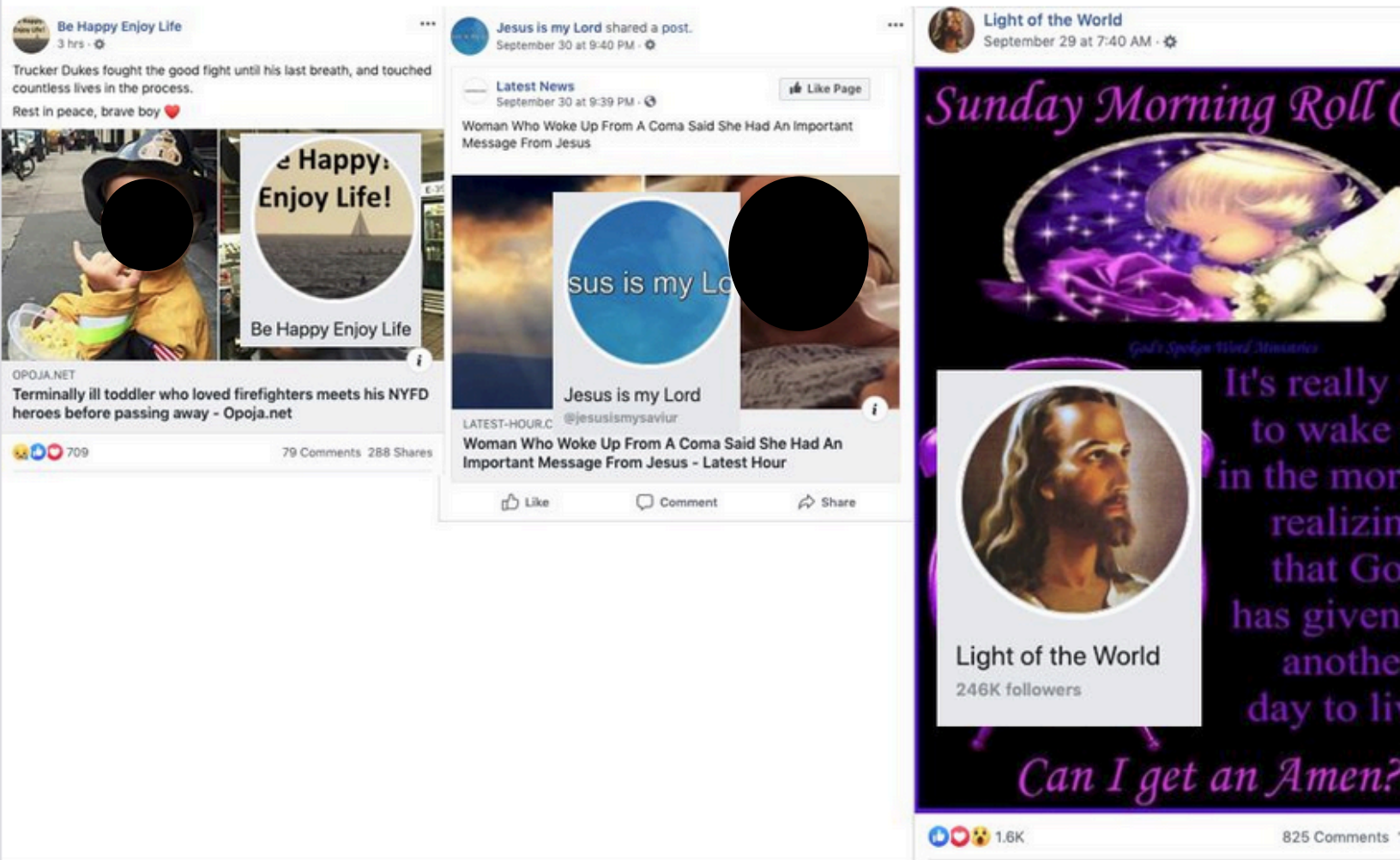
Ranked Troll Farm + Authentic Pages

name	is_mkxx_page	page_id
OVERALL_MKXX	1	0
Be Happy Enjoy Life	1	249346019035747
Jesus is my Lord	1	944383662417356
You Can	1	746720989059699
Speak Loudly	1	337270806881846
Purpose of Life	1	175964166645416
Positive Quotes	1	1537484389716090
Memories of a loved one	1	457990627606884
Smile & Shine	1	658641567851336
Light of the World	1	171049673814368
Why Not Us?	1	190698914861953
Happiness Is In Your Mind	1	1974991652809180
My Life	1	749005765476444
Life is Beautiful	1	277543839631924
Blessing	1	291773748400681
Conscious Reminder	1	160340534382111
Blessed	1	437169703776552
Guideposts	0	78775325627
Simply Beautiful	1	272199883444219
Motivational Quotes & Videos	1	373989519691000
light-candle for a lost love one	1	45298736196
I Believe In God Do You?	1	999959740142056
Christianity Today	0	16865030420
Life Church	0	5898823795

On the left: The top Troll Farm Pages built around content from religious communities, ranked by reach. On the right: Similar to the left, but the top authentic religious Pages are included for comparison.

The top Pages are overwhelmingly Troll Farms. The first authentic Page is 15 positions down. And the next after are 20.

Again, the content isn't violating or borderline. It does skew very hard towards clickbait and engagement bait, however.



Example posts from Troll Farm Pages built around religious content.

The “Trucker Dukes fought the good fight until his last breath...” story is a prime example of the very strange path stories can take before they land on our platform.

The Trucker Dukes story actually begins on a FB Page, “TeamTrucker”. From there, it was picked up by The Daily Mail in 2017. For some reason, two years later, the Epoch Times decided to re-write the Daily Mail story, using a technique called article spinning [redacted] “The Fine Art Of Article Spinning”. The article was spun again by Newsner the day after Epoch Times published it. It was scraped from Newsner by Opoja.net, which presumably is run by the Troll Farmers. And today, they posted it to their FB Page.

So, the path this story took was...

Authentic, genuinely beautiful and touching FB Page 3 years ago → Daily Mail article 2 years ago → Epoch Times article a month ago → Newsner article a day later → Opoja[.]net article promptly after → Troll Farm run FB Page today.

It is sad that this type of content ecosystem is rewarded by our platform. The religious Pages tend to be a mixture of the heavy engagement bait like the meme from “Light of the World” and sensational click bait stories about children, animals, and police officers dying, like the “Be Happy Enjoy Life” and “Jesus is my Lord” example.

It's also strange that the Feed ranking algorithm, which is supposed to be assessing the relevance of content to our users, is deciding that 2 year old tabloid stories are relevant to our users. It is a weird editorial stance to take, that multiyear old tabloid stories are still relevant.

Native American Pages

Here are the top Troll Farm Pages targeting Native American communities on FB. Or people interested in the Native American community. Top authentic Pages sourced from [here](#).

Ranked Troll Farm Pages

name	is_mkxxk_page	page_id
OVERALL_MKXXK	1	0
Native Americans	1	481981788887455
Native Americans Cherokee	1	328723524658639
Eagle Spirit	1	395613751010806
Bald Eagles spirit	1	302562687038584

Ranked Troll Farm + Authentic Pages

name	is_mkxxk_page	page_id
Indian Country Today	0	3761908474
OVERALL_MKXXK	1	0
American Indian College Fund	0	1210331493
Association on American Indian Affairs	0	2163687105
Native Americans	1	4819817888
Native Americans Cherokee	1	3287235246
Native American Times	0	1116963902
Native American Rights Fund	0	1337405595
National Congress of American Indians	0	2445547496
Redhawk Native American Arts Council	0	1090241619
Eagle Spirit	1	3956137510
Bald Eagles spirit	1	3025626870
Native American Journalists Association	0	2361736964

On the left: The top Troll Farm Pages built around content from Native American communities, ranked by reach. On the right: Similar to the left, but the top authentic Native American Pages are included for comparison.

The content they typically post is a mixture of engagement bait, pictures of people and children, and politically oriented memes.



Example posts from Troll Farm Pages built around Native American content.

How They Operate

From these examples, a general picture emerges. The Troll Farm owners follow creators in a community that produce engaging content, or communities where it gets posted, and they scrape and aggregate the most engaging content. They then post it to their Pages, and watch the content “go viral” and gain a lot of distribution.

They don’t create original content for a couple reasons. Primary is, it is way easier to take others content than to create your own. Second is, they do not have any knowledge of the communities they target and wouldn’t know how to make relevant content. How could they make a BlackBabies Page, when they are thousands of miles away from any African American babies? [Buzzfeed has even interviewed some of them](#) about their production tactics, and they admit to stealing all their content.

The Pages are all totally anonymous, because they are not running these Pages authentically. They have little interest in building any connection with their audience. They probably have very little respect for their audience. We have testimonials from people who engage in this activity [redacted] [“Exploiting Communities”](#)], and they express disdain for their followers.

How Has This Happened?

So, the largest religious oriented Pages are run by non-religious people, the largest African American oriented Page is run by people who probably have never met one, and inauthentic Pages seem to largely be winning in our public content distribution channels. How did we get here?

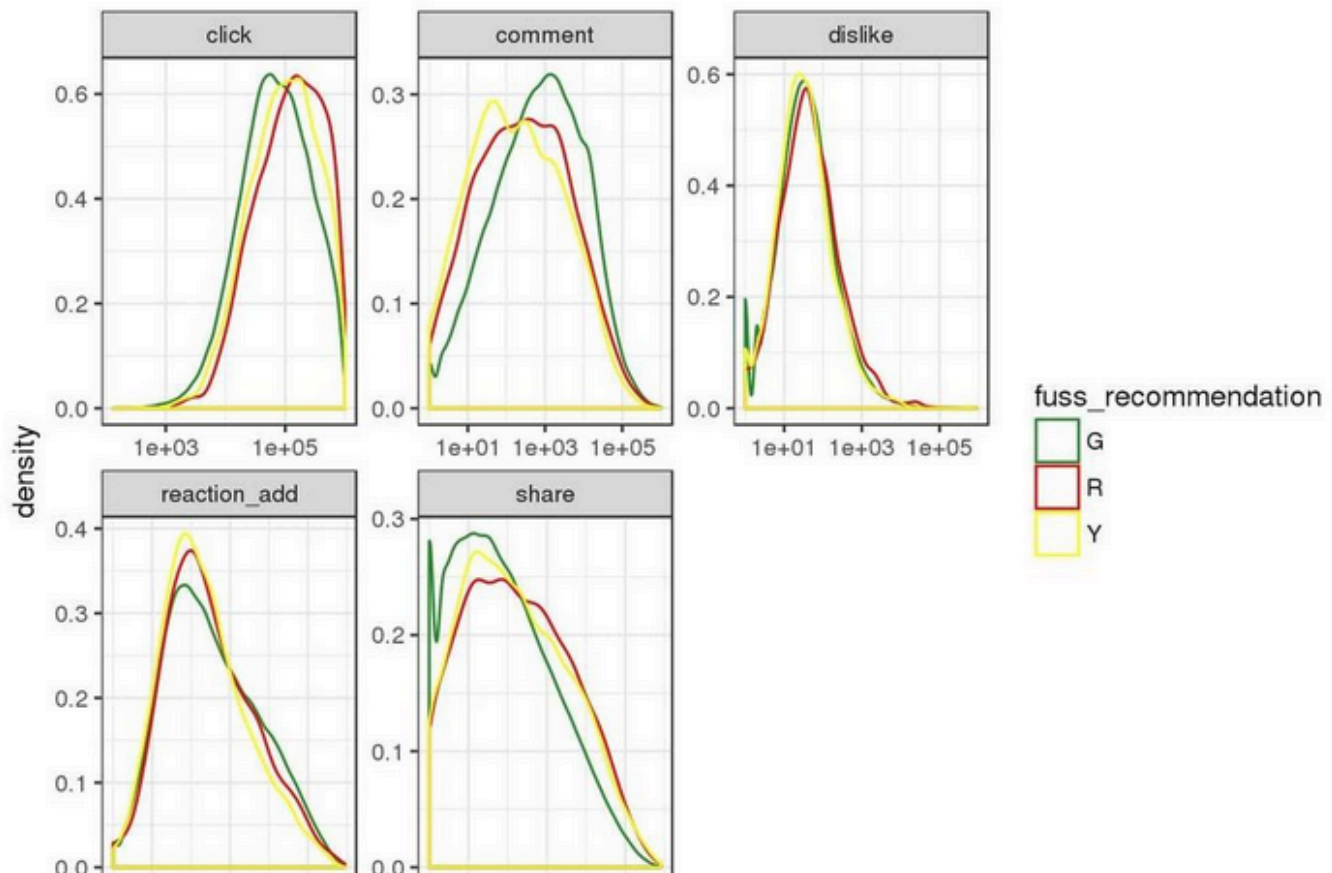
It is important to note that 75% of their reach is non-direct. *This means that a big majority of their ability to reach our users comes from the structure of our platform and our ranking algorithms, rather than user choice.* Instead of users choosing to receive content from these actors, it is our platform that is choosing to give them an enormous reach.

There are three fundamental properties of our platform that enable them.

1. Easily exploited engagement based ranking systems
2. A lack of basic media literacy checks to influence ranking
3. Edge stories which expand reach beyond direct fans

Engagement based ranking is easily exploited

At the heart of Feed ranking, there are models that predict the probability a user will take an engagement action. These are colloquially known as $p(\text{like})$, $p(\text{comment})$, and $p(\text{share})$ etc. These systems have been studied from an integrity angle several times, and it has been shown that content that scores highest in these fundamental models heavily skews towards content we know to be bad through FUSS [██████████, "FUSS, Feed Ranking, and Gradient Descent"].

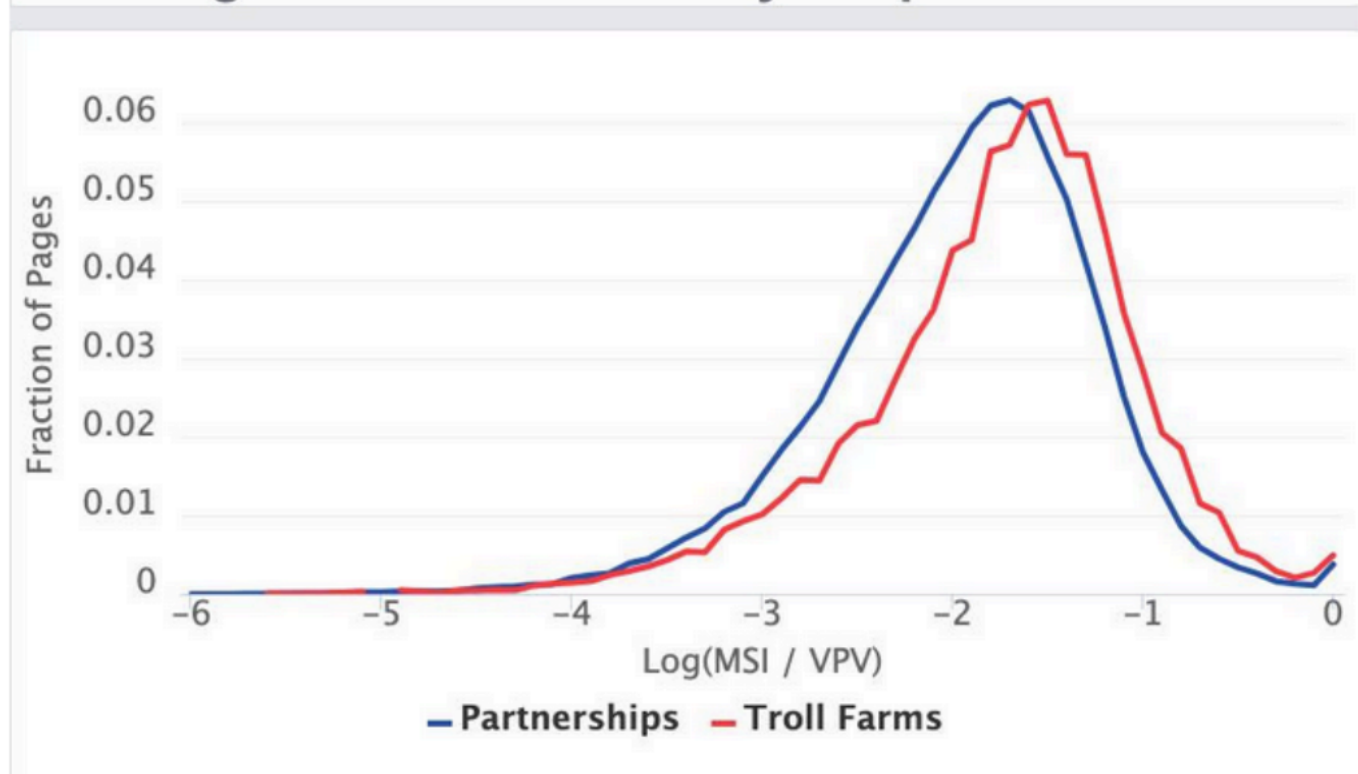


The distribution of $p(\text{click})$, $p(\text{comment})$, $p(\text{dislike})$, $p(\text{reaction})$, $p(\text{share})$ for FUSS G, Y, and R content. For click, share, and dislike, R and Y content wins at the high end. Image Credit [REDACTED] used without permission cause Former People Of FB Issues. [https://fb.workplace.com/notes/\[REDACTED\]/fuss-feed-ranking-and-gradient-descent/1982672198728761/](https://fb.workplace.com/notes/[REDACTED]/fuss-feed-ranking-and-gradient-descent/1982672198728761/)

In addition, bad actors have learned how to easily exploit the systems. Basically, whatever score a piece of content got in the models when it was originally posted, it will likely get a similar score the second time it is posted. Because all social media platforms provide engagement signals on content, bad actors can scrape all the social media platforms, harvesting the most engaging content within various communities, and repost it to FB to watch it go viral all over again [REDACTED], "On Manufactured Virality"].

You can see their success when comparing their average MSI per VPV to that of our partnership Pages.

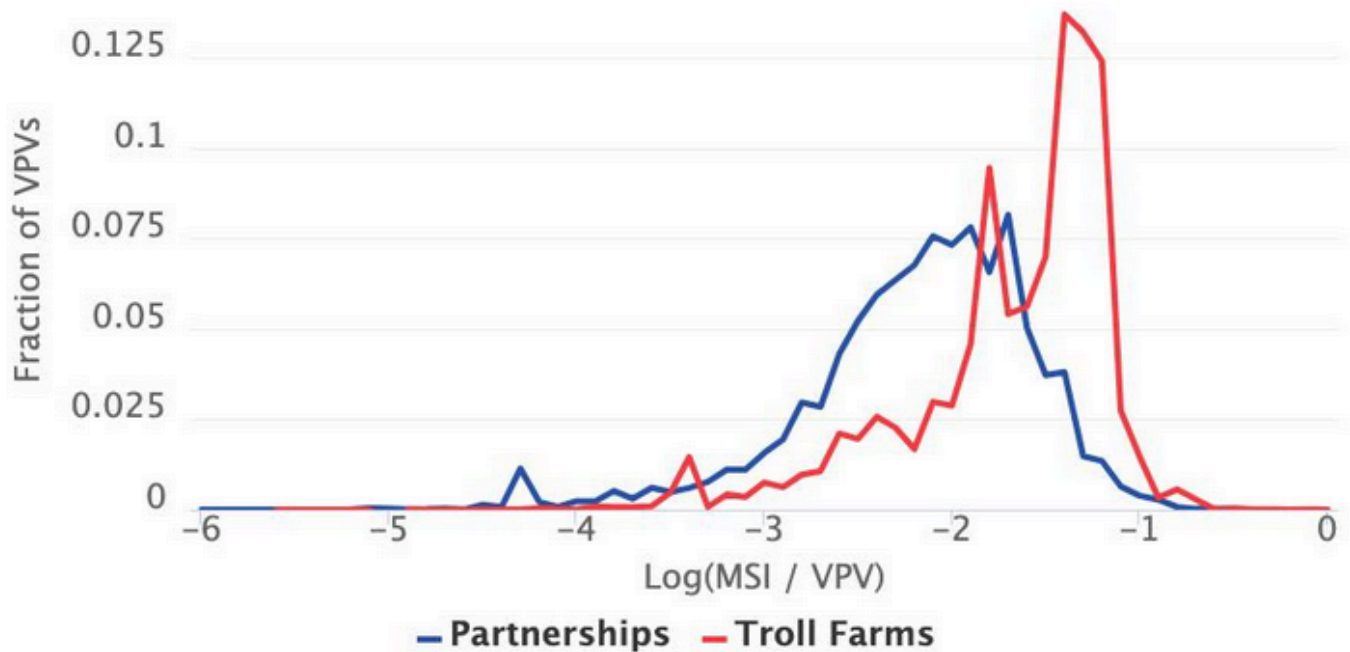
Page Count Distribution by MSI per VPV Ratio



Distribution of MSI per VPV ratios for Partnership Pages and Troll Farm Pages. Troll Farm Pages score higher on MSI per VPV. MSI is a fundamental thing we try to optimize in Feed ranking. So more VPVs should be given to Troll Farms here.

At the highest end of the MSI per VPV ratios, the Troll Farms are more frequent relative to other Pages. Because our system is largely optimized to maximize MSI, this translates into a large distribution boost. Here is the same distribution but weighted by impressions.

Page VPV Distribution by MSI per VPV Ratio



Similar to previous plot, but now each Page is weighted by VPVs. This shows that the edge in MSI per VPV that Troll Farm Pages have translates to a substantial VPV boost.

The extra edge in the MSI per VPV ratio translates to a large increase in distribution for the Troll Farms.

Lack of Basic Media Literacy Checks

Media Literacy comes down to [asking a set of basic questions about any media consumed](#). Such as “Who created this content?” “How did they create this content?” “What did they do to amplify this content?” and “Why did they create this content?”.

Troll Farm content is always totally anonymous, and the Troll Farms exclusively upload content that they did not create. So they get a total failure for the first two checks of basic media literacy.

But this doesn't matter. Because the impact this has on our ranking and distribution systems is negligible. So there is nothing preventing a totally anonymous FB Page that posts exclusively content they did not create from building a large audience.

Edge Stories Expanding Reach

On FB, a user who doesn't follow a Page might be exposed to the content through and edge story. If a friend of theirs comments or reshares the content. So, if a Page posts content that has a high comment or reshare rate, the Page will be able to reach an audience well outside its followers.

The Troll Farms combine these three facts into a potent attack on our platform. They aggregate content that has a high comment and reshare rate, which skews towards clickbait, borderline, divisive, and sensational content due to our models, and post at high rates to reach well beyond their follow graph and build a large audience. All while not being transparent with their audience or creating a single piece of original content.

The ease of exploiting our ranking systems is widely known and discussed externally [REDACTED] ["Comprehensive Guide To Becoming A Video Publisher On FB"](#)]. We should not be surprised that bad actors have become experts at it. There are some workstreams that hope to address this [REDACTED] ["Unoriginal Content on Civic and Business Pages is High-Prevalence"](#)].

Why Do They Do It?

Money is definitely a big reason. There are a few ways to monetize an FB page.

1. Link to a domain where you show ads
2. Enter into our partnership programs
 1. Instant Articles
 2. Ad Breaks
3. Do sponsored posting and sharing
4. Sell the Page after it reaches a large size

They engage in all of these activities.

Linking to a domain is the easiest, and you can see them doing it in the religious Pages. Two is the most worrisome for us.

Troll Farms In Our Partnership Programs

Troll Farms have been able to get into both Instant Articles as well as Ad Breaks. Fortunately in H1, a lot of work was done to get them out [REDACTED], ["Update: Unoriginal Content Enforcement and corresponding video upload activity"](#)].

But in general, content producers with similar practices still have non-zero prevalence in our partnership programs. In Instant Articles, there was a period when perhaps as much as 60% of Instant Article reads were taking place on scraped content, which is the Troll Farms article writing method of choice [REDACTED], ["Instant Articles content originality analysis"](#)].

Our partnership programs are in general very appealing to the Troll Farms, and so there will always have to be a lot of effort and vigilance to keep them out.

Are There Any Regulation Or Policies That Apply To Them?

Absolutely. IP violations are crimes, and although we don't have 100% certainty that the Troll Farms don't have rights to the content they post... we can be pretty confident. I'm obviously not a lawyer, so can't give any actual legal interpretation here. But in my experience, the legal situation here is a mess.

IP protection on the internet currently doesn't really work. For us to remove a piece of content, we have to wait for the owner to discover the IP violation, give us a DMCA notice, and then we must review the content and action it. This process takes time. And there have been numerous complaints by content creators that in the time window it takes for them to find their IP on our platform, report it, and have us action on it, the stolen content has already amassed the vast majority of distribution it would ever receive on our platform. This has caused a lot of strife between Facebook and the Youtube creator community, for example [REDACTED].

Unoriginal content is in fact against our policies. We have a Pages policy in place that [forbids Pages from posting content that has limited originality](#). And it is [against our Monetization Eligibility Standards as well](#). We currently do not enforce our own policies at the Page level, even towards known bad actors. We need to get way better at enforcing these.

The current state of Safe Harbor law has forced us into this weird situation where we have to be very careful about proactive identification of unoriginal content, so that we don't slip into a territory where we are proactively making IP assessments. Because then we open ourselves to litigation. It's a bizarre situation. And the Troll Farms are currently greatly benefiting from it. It would be awesome to see FB working with lawmakers to find a better solution here.

But the bottom line is, regardless of if there is or is not IP violation taking place, posting strictly unoriginal content violates our policies and exposes our communities to exploitation. So we should beef up our capabilities here.

And, sometime very soon, we should free ourselves from the Safe Harbor constraints around proactive identification of IP violations. Youtube does proactive identification like crazy. It is irresponsible for a platform of our size to be hiding behind Safe Harbor. We should be champions of content creators.

How Many Actors Are Similar To The Troll Farms?

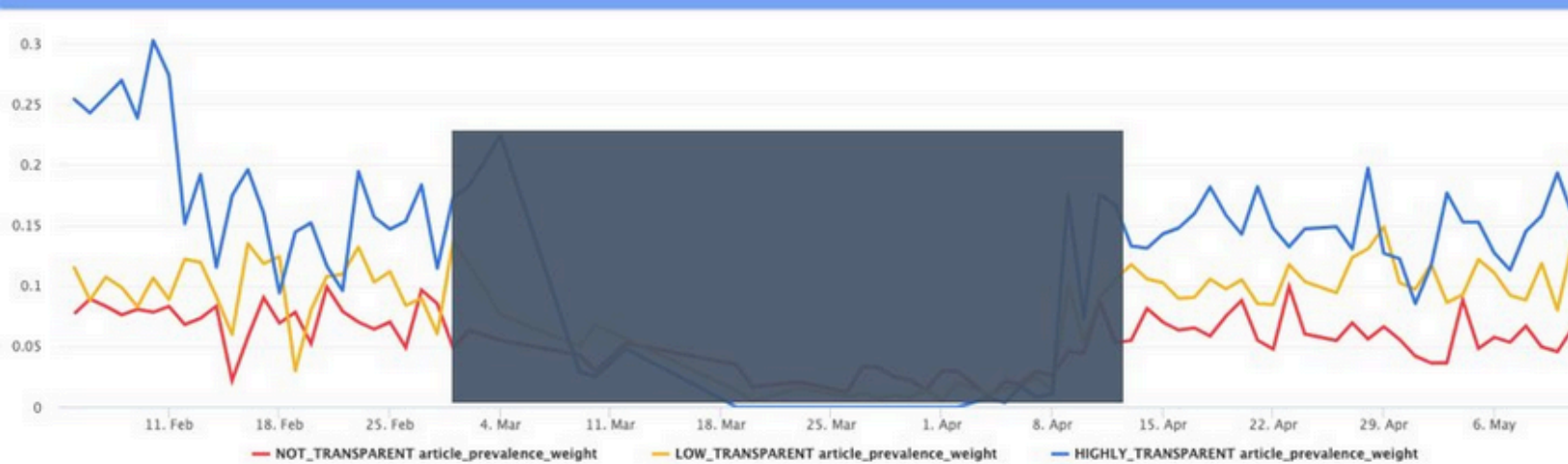
If we ask how many Pages are engaging in similar behavior to the Troll Farms (Exploiting our ranking systems and distribution channels and not passing basic media literacy checks), we find that the prevalence is substantial.

In one of the earlier Pages ecosystem analysis, a rough estimate suggested that as much as 40% of US Page VPs could be going to Pages that had transparency issues and posted primarily unoriginal, or limited originality, content. [REDACTED] ["The Pages Ecosystem And Exploitative Practice Prevalence"](#)]

There were also studies specifically within the video space. There, it was seen that around 60% of video watch time was going to video aggregators. [REDACTED] [“What have we learned from unoriginal content demotion?”](#)]

In the transparency domain, we measured the prevalence of actors with low and high levels of transparency. The prevalence of Pages with serious transparency issues was between 15% and 20% [[“Low Quality Rater Workflow”](#)]

Transparency Bucket Prevalence



The prevalence of Pages with High, Low, and No transparency, in terms of US Pages VPVs. Pages with No or Low transparency would fail to provide users with any meaningful details, and fail media literacy checks. Together they are between 15% and 20% of all US Page VPVs.

You can also look to see the prevalence of Pages that are exploiting our virality channels. An estimate of the prevalence of “manufactured virality”, which is repeatedly posting content that went viral before or on other platforms, suggests that up to 50% of all reshares impressions could be going to actors abusing the reshare mechanism of our platform [REDACTED], [“On Manufactured Virality”](#)]. Which equates to around 25% of all Pages VPVs.

This all adds up to a significant fraction of all Page VPVs. Likely well into double digits.

This should not be surprising. The truth is that this is the best strategy to create a large Page on FB! If a friend came to me and said they wanted a really big FB Page, and didn't particularly care about anything than the fan count, I would tell them to run the Troll Farm playbook. This playbook has been discovered by many actors all over the world, and many are running is successfully today. The Troll Farms from the Balkans are a bit under 1% of Page VPVs, but they are much less than a tenth of the overall problem.

What Do Our Users Think Of These Pages?

Our users sit in two camps with these Pages.

One camp doesn't realize the Pages are run by inauthentic actors who are exploiting their communities. They tend to love these Pages. They like how entertaining the posts are, and how they reaffirm their already held beliefs.

The other camp does realize the Pages are run by inauthentic actors. They hate the every loving shit out of these Pages. They hate these Pages with a passion that even I find impossible to match.

Here is a real quote from a real user who discovered the Page was run by an inauthentic network of foreign admins.

This is a Fake page ran by Non Indians out of multiple locations overseas , They have a network of fake pages and constantly steal art and resale it on t shirts and other items , but most don't get there product , it's a large Scam , this whole group is a fraud ring that also has there hands in real-estate and other criminal ventures , these pages are truly Culture vultures , and it's not just with the native community , they exploit people from all over the world for profit

I hate these Pages. But I can't match that. This review, by the way, is where I got the phrase "Exploiting Community" from. So our users are literally trying to tell us that they feel exploited by these Pages.

When a user understands the Page is inauthentic, the second most common complaint is the stolen content. Users will sometimes recognize content coming from these Pages because they happen to also follow the original uploader. They frequently complain about the content theft, because they are trying to help and protect a content creator in their community that they value [REDACTED] "[Exploiting Communities With The Troll Farms](#)"]

What Can We Do About It?

The mission we have taken on as employees of this company is to empower people to build community and bring the world closer together. Within the public content space, our primary tool in making progress towards that mission is the feed ranking of public content.

When the content producers that win that system are exploiting communities on our platform rather than building and supporting them, it becomes clear that the ranking system does not reflect our company values. So much so that it is actually working against us.

The basic tactics of the Troll Farms has changed very little since 2016 when they were first uncovered. They discuss politics less. So that is an improvement (Although they do discuss it at a non-0 level). Their off platform monetization has been made more difficult, but the operations are still highly profitable. Basically, we have yet to implement a comprehensive solution to reduce their overall success.

But, the good news is that this is far from the first time humanity has fought bad actors in our media ecosystems. And it is even far from the first time web platforms have fought similar bad actors. There is a proven strategy to aligning media ecosystems and distribution platforms with important missions, such as ours, and societal value.

The successful strategies demonstrated by Google Search, Google News, and to a lesser extent Youtube, provide action items at both the short and long term we could pursue.

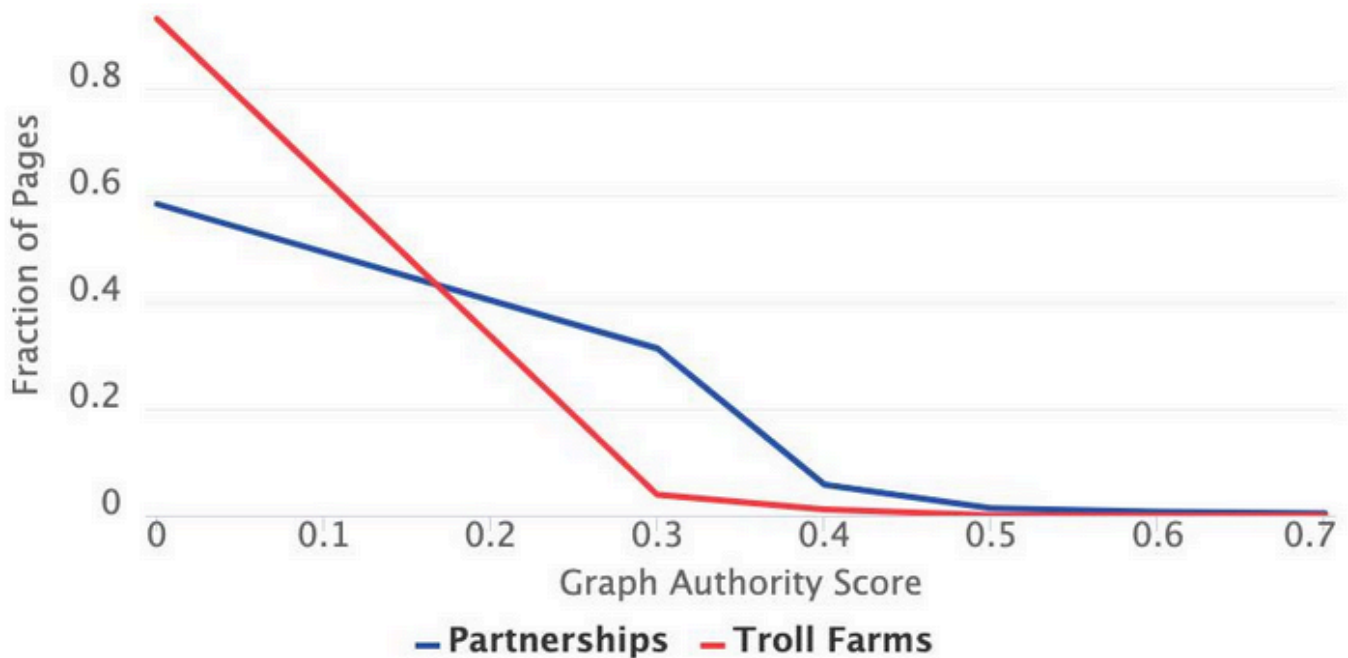
Google In The Short Term

Google Search uses a ton of varied and creative features in determining the ranking of results. Their most important is now so old that its patent has expired. PageRank is basically the algorithm that built the internet. From 2000 to ~2012 at least. And it's actually a much older idea. Graph based authority measure date back to the 70's.

Domain based PageRank signals are still, to this day, [the most important ranking features in Google Search](#). And we have our own implementation of a graph based authority measure [██████████ ["Introducing GraphAuthority Inside FB"](#)]. Which we can use to directly calculate the Graph Authority of Pages.

Earlier we looked at the distribution of MSI to VPV ratios for partnership Pages and Troll Farm Pages. And the Troll Farm Pages were clearly winning in that metric. Here's the distribution for Graph Authority.

Page Count Distribution by Graph Authority Score



Distribution of Graph Authority for Partnership Pages and Troll Farm Pages. Partnerships score higher. If Graph Authority was an important signal in ranking, then it would reduce the prevalence of the Troll Farm Pages.

Graph Authority flips the trend. 90% of the Troll Farm Pages have exactly 0 Graph Authority, which is not surprising, because we would not expect external parties to vouch for them. Partnerships clearly win here.

Our current Feed ranking models know almost nothing meaningful about content producers. Beyond what groups of users tend to engage with the various groups the Page is in. It's a lot of extremely fancy collaborative filtering algorithms, but all based around engagement.

Adding even just some easy features like Graph Authority, and pulling the dial back from pure engagement based features, would likely pay off a ton in both the integrity space, and long term likely in engagement as well.

The Transition From Yellow Journalism To Professional Journalism

In the long term...

Ok, this is way back. But. In the mid to late 1800's in America, newspapers were sold by young boys standing on the street corner shouting things. There were many different newspapers, and none of them were particularly old or established. So people didn't have any real loyalty or trust in any of them.

In this situation, where each newspaper is essentially unknown to its readers, what should the young boys shout?

Obviously, sensational clickbait stories. And misinfo. And stories that stoked fears and reinforced existing biases. When the content producers are invisible, and only the content matters, and the sales of yesterday don't predict the sales of tomorrow, there is little other choice! This was the incentive structure at play that gave us the era of yellow journalism.

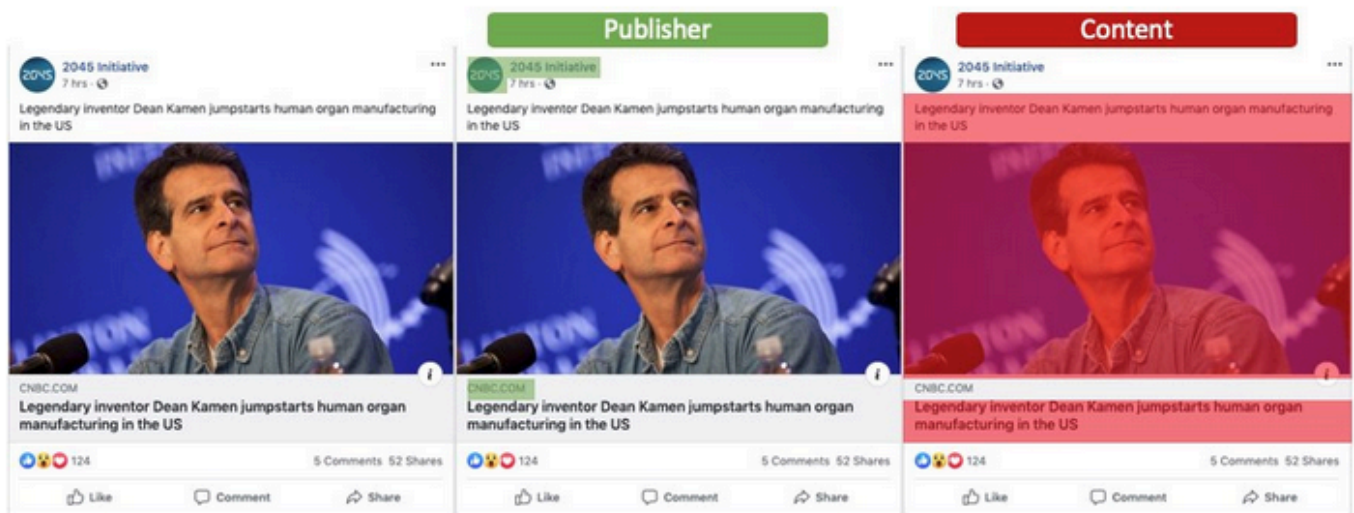
This changed between 1890 and 1910. There were two big shifts. The first was, having such a polluted ecosystem was causing societal harm, and people were getting pretty sick of it. So there was public backlash against the media at large. The second shift came from Adolph Ochs when he took over the New York Times. His idea was simple but revolutionary. What if, instead of having the boys shout crazy stories, you could actually build a brand, and the boys would simply shout the brand name. Instead of "Elderly couple mauled by rabid dogs, read all about it!" what if they said "Get your New York Times here, all the news that's fit to print!"

The strategy Ochs took was to capitalize on the backlash against sensationalized journalism. He instituted practices around reporting, an aspiration of always being truthful. All that good stuff.

In the early 1900's, the New York Times was experiencing Facebook levels of year over year growth. It worked. In 1914, the AP first set out its code of ethics. And Journalism became a respected profession within society.

Hopefully some of this is resonating. If you want the full history, read "Discovering the News" by Michael Schudson. It's great.

It's kind of amazing how much of this is just visible in how we present content to our users.



The number of pixels dedicated to publisher information and branding in our posts is tiny compared to the content. And all posts are formatted identically. This means publishers likely have trouble establishing recognition with their audience, and basically all publishers are competing at the content level.

I'm sure I'm the 50 billionth person to point this out. But like... with the raw pixel counts going to publishers and content, it's no wonder that the publishing world is resorting to sensationalism, clickbait, misinfo. It's no wonder "newspapers" and "publishers" have become disposable shells that people build and discard at whim.

It's no wonder foreign actors have no problem inserting themselves into our national discourse for narrow narrow reasons.

I have no problem with Macedonians reaching US audiences. I'm totally down with it! I'd love to see what amazing culture they have there. I'd love to hear what they think of ours. And hey, Mother Teresa was born in Skopje, now the capital of Macedonia. If Mother Teresa had wanted to use FB to deliver a message to America? I think that'd have been wonderful.

But if you just want to write python scripts that scrape social media and anonymously regurgitate content into communities while siphoning off some monetary or influence reward for yourself... well you can fuck right off. I don't care how many sessions you generate. I don't know why anyone gets excited about those sessions.

So Let's Just Follow Googles Lead

Google has set a remarkable example of how to build Ochs' idea into a web platform. How to encode company values and missions into ranking systems. Figuring out how to make some of it work for FB and IG would provide the whole company with enormous value.

Google is remarkably transparent about how they work and how they fight these types of actors. If you haven't read "[How Search Works](#)" I highly recommend it. It is an amazing lesson in how to build a world class information retrieval system. And if you haven't read "[How Google Fights Disinformation](#)", legitimately stop what you're doing right now and read it.

The problem of information retrieval (And Newsfeed is 100% an information retrieval system) comes down to creating a meaningful definition of both the quality of the content producer and the relevance of the content. Google's basic method was to use their company mission to define the quality.

Google's mission statement is to make the world's information widely available and useful. The most important word in that mission statement is "useful". A high quality content producer should be in alignment with the IR systems mission. In the case of Google, that means a content producer that makes useful content. A low quality producer makes content that isn't useful. Google has built a completely objective and defensible definition of what useful content is that they can apply at scale. This is done in their "[Search Quality Rater Guidelines](#)", which they publish publicly.

The way Google breaks down the utility of content basically lands in 3 buckets. How much expertise does the author have in the subject matter of the content, as determined by the credentials the author presents to the users. How much effort does the author put into their content. And the level of 3rd party validation the author has.

If the author has 0 experience in the subject, doesn't spend any time on the content, and doesn't have any 3rd party validation, then that author is going to be labeled lowest quality by Google and hardly get any search traffic. Does that description sound familiar? It is a pretty solid description of the Troll Farms.

Google calls their quality work their first line of defense against disinformation and misinformation. All we have to do are figure out what the objective and defensible criteria are for a Page to build community, and bring the world closer together. We are leaving a huge obvious win on the table by not pursuing this strategy.

If you want to see what that strategy could look like practically, there is [REDACTED] "[Making Web Platforms Resilient](#)"] and [REDACTED] "[Mission Alignment and Comprehensive Authenticity](#)"].

Conclusion

It will always strike me as profoundly weird that the largest Page on FB posting African American content is... run out of Kosovo. That's so weird! And genuinely horrifying.

It seems quite clear that until that situation can be fixed, we will always be feeling serious headwinds in trying to accomplish our mission. Newsfeed and specifically ranking is such an integral part of our platform. For almost everything we want to accomplish, Feed plays a key role. Feed is essential enough that it doesn't particularly need any mission beyond our companies. FB, and IG, need to figure out what implications our company mission has on ordering posts from users inventory.

Until we do, we should expect our platform to continue to empower actors who are antithetical to the company mission.

Signing Off

I'm sad that I'm leaving FB and won't be able to work on this problem with the level of impact that FB provides. But I am heartened by the amazing people who have previously and will continue to try and reduce the success of these actors.

[REDACTED] "[What are the Copy-Pasta'ers up to Now?](#)", "[Hoax Page Archetypes](#)"]
[REDACTED] "[3 People - 25% of US Civic Misinfo](#)" "[Yet another foreign-admined LIS](#)"]